Conference Abstract

IndexMEED cases studies using "Omics" data with graph theory

Romain David[‡], Jean-Pierre Féral[‡], Anne-Sophie Archambeau[§], Fanny Arnaud[|], David Auber[¶], Nicolas Bailly[#], Loup Bernard[¤], Laure Berti-Equille[«], Cyrille Blanpain[»], Vincent Breton[ˆ], Anne Chenuil-Maurel[ˇ], Anna Cohen Nabeiro[¦], Alrick Dias[ʔ], Aurélie Delavaud[ˤ], Robin Goffaud[¢], Sophie Gachet[ˇ], Karina Gibert[ℓ], Manuel Herrera Fernandez[ˀ], Luc Hogie[P], Dino Ienco[Ꜹ], Romain Julliard[ℇ], Yvan Le Bras[F], Julien Lecubin[Ꞙ], Yannick Legre[Ɲ], Michelle Leydet[Ʞ], Grégoire Lois[Ɠ], Bénédicte Madon[ʔ], François Marchal[₩], Victor Mendez Munoz[Ꞇ], Jean-Charles Meunier[‡‡], Jean-Baptiste Mihoub[F], Isabelle Mougenot[§§], Sophie Pamerlon[||], Eric Peletier[¶¶], Geneviève Romier[##], Dad Roux-Michollet[¤¤], Alison Specht[««], Christian Surace[»»], Jean-Claude Raynal[ˆˆ], Thierry Tatoni[ˇ]

‡ Mediterranean Institute of Biodiversity and marine and terrestrial Ecology (IMBE), Aix Marseille Université/CNRS/IRD/ Université d'Avignon, Station Marine d'Endoume, Marseille, France, Metropolitan
§ GBIF France, Paris, France, Metropolitan
| ENS Lyon, OHM Vallée du Rhône, Lyon, France, Metropolitan
¶ LABRI, Bordeaux, France, Metropolitan
# Hellenic Centre for Marine Research (HCMR), Gouves, Greece
¤ ARCHIMEDE-UMR 7044, Université de Strasbourg CNRS, Strasbourg, France, Metropolitan
« IRD (ESPACE DEV U228) and LIF, Marseille, France, Metropolitan
» SIP OSU Pytheas, CNRS, Marseille, France, Metropolitan
ˆ IdGC – LPC, CNRS, France Grilles, Marseille, France, Metropolitan
ˇ Mediterranean Institute of Biodiversity and marine and terrestrial Ecology (IMBE), Aix Marseille Université/CNRS/IRD/ Université d'Avignon,, Marseille, France, Metropolitan
¦ ECOSCOPE, FRB, Paris, France, Metropolitan
ʔ Mediterranean Institute of Biodiversity and marine and terrestrial Ecology (IMBE), Aix Marseille Université/CNRS/IRD/ Université d'Avignon,, Marselle, France, Metropolitan
ˤ ECOSCOPE, FRB, Marseille, France
¢ ECOSCOPE, FRB, Marseille, France, Metropolitan
ℓ Department of Statistics and Operations Research, Universitat Politecnica de Catalunya, Barcelona, Spain
ˀ EDEn - Dept. of Architecture and Civil Eng., University of Bath, Bath, United Kingdom
P I3S (the laboratory of Computer Science of the University of Nice-Sophia Antipolis) and Inria, Nice, France, Metropolitan
Ꜹ UMR TETIS, Montpellier, France, Metropolitan
ℇ Museum of Natural History (MNHN), Paris, France
F CESCO - Centre d'Écologie et des Sciences de la Conservation Muséum national d'Histoire naturelle, Paris, France
Ꞙ SIP OSU Pytheas, Marseille, France, Metropolitan
Ɲ European Grill Infrastructure, Amsterdam, Netherlands
Ʞ Mediterranean Institute of Biodiversity and marine and terrestrial Ecology (IMBE), Aix Marseille Université/CNRS/IRD/ Université d'Avignon, Marseille, France, Metropolitan
Ɠ CESCO - Centre d'Écologie et des Sciences de la Conservation Muséum national d'Histoire naturelle, Paris, France, Metropolitan
ʔ Université Bretagne Occidentale IUEM, Brest, France, Metropolitan
₩ UMR 7268 ADES - Anthropologie Bioculturelle, Droit, Ethique et Santé Université d'Aix-Marseille / CNRS / EFS, Marseille, France, Metropolitan
Ꞇ Department of Computer Architecture and Operating Systems (CAOS) Universitat Autònoma de Barcelona (UAB), Barcelona, Spain
‡‡ LAM / CeSAM, Marseille, France, Metropolitan
§§ UMR Espace DEV, Montpellier, Montpelier, France, Metropolitan

|| GBIF France, Paris, France

¶¶ Institut de Génomique - Genoscope - CEA, Paris, France, Metropolitan

IdGC, CNRS, France Grilles, Lyon, France, Metropolitan

¤¤ GRAIE, OHM Vallée du Rhône, Lyon, France, Metropolitan

«« CESAB, FRB, Aix en Provence, France, Metropolitan

»» LAM, CNRS, Aix-Marseille Université, Marseille, France, Metropolitan

^^ ECCOREV FR3098, CNRS, Aix-Marseille Université, Marseille, France, Metropolitan

Corresponding author: Romain David (romain.david@imbe.fr)

Abstract

Data produced within marine and terrestrial biodiversity research projects that evaluate and monitor Good Environmental Status, have a high potential for use by stakeholders involved in environmental management. However, environmental data, especially in ecology, are not readily accessible to various users. The specific scientific goals and the logics of project organization and information gathering lead to a decentralized data distribution. In such a heterogeneous system across different organizations and data formats, it is difficult to efficiently harmonize the outputs. Few tools are available to assist. For instance standards and specific protocols can be applied to interconnect databases. Such semantic approaches greatly increase data interoperability.

This communication present the recent results and the consortium IndexMEED (Indexing for Mining Ecological and Environmental Data) activity that aims to build new approaches to investigate complex research questions, and support the emergence of new scientific hypotheses based on graph theory Auber et al. 2014). Current developments in data mining based on graphs, as well as the potential for relevant contributions to environmental research, particularly about strategic decision-making, and new ways of organizing data will be presented (David et al. 2015). In particular, the consortium makes decisions on how i) to analyze heterogeneous distributed data spread throughout different databases combining molecular and habitat characteristics data [3], ii) to create matches and incorporate some approximations, iii) to identify statistical relationships between observed data and the emergence of contextual patterns using a calculation library and distributed calculation center at the European level, iv) to encourage openness and sharing data while complying with the general principles of FAIR (Findable, Accessible, Interoperable, Reusable and citable) in order to enhance data value and their utilization. IndexMEED participants are now exploring the ability of two scientific communities (ecology sensu lato and computer sciences) to work together, using several studies cases. The ECOSCOPE

project aims to meet the need to access structured and complementary omics-datasets to better understand biodiversity state and its dynamics. Indeed, the ECOSCOPE case study targets to visualize, through the graph approach, links between datasets and databases from genetics to ecosystems. Another case study, displaying anthropology fossils and omics on the same graph, will also be presented. DEVOTES (DEVelopment Of innovative Tools for understanding marine biodiversity and assessing good Environmental Status) and CIGESMED (Coralligenous based Indicators to evaluate and monitor the "Good Environmental Status" of the MEDiterranean coastal water) European projects, conducted by IMBE, are focused on photo quadrats, cartography and omics data of the marine hard bottom in order to discover context patterns helpful to build decision support system building. Study case "65 Millions d'observateurs" French project is testing AskOmics to provide a graph-based querying interface using RDF (Resource Description Framework) and SPARQL technologies.

Scientific questions can be resolved by the new data mining approaches that offer new ways to investigate heterogeneous environmental data with graph mining (Muñoz et al. 2017). The uses of data from biodiversity research demonstrate the prototype functionalities (*David et al. 2016*) and introduce new perspectives to analyze environmental and societal responses including decision-making at large scale, both at the information system level and the observing system level than at the observed system level.

Keywords

Interdisciplinarity, Data qualification, Omics data, Graph, Thesaurus, Decision Support Tools

Presenting author

Romain David

Funding program

Labex DRIIHM (OHM Bassin Minier de Provence, OHM Vallée du Rhône, OHM Littoral méditerranéen), Fédération ECCOREV FR 3098, OSU Pythéas, and LabEx OT Med

Hosting institution

CESAB, ECOSCOPE, FRB, GBIF, IMBE, LAM,

References

- Auber D, Archambault D, Bourqui R, Delest M, Dubois J, Pinaud B, Lambert A, Mary P, Mathiaut M, Melancon G (2014) Tulip III. Encyclopedia of Social Network Analysis and Mining. https://doi.org/10.1007/978-1-4614-6170-8_315
- David R, Tatoni T, Féral J, Dias A, Lecubin J, Blanpain C, Surace C, IndexMed lc (2016) Bilan Juin 2015 – Février 2016 du Projet VIGI-GEEK : VIsualisation of Graph In trans-disciplinary Global Ecology, Economy and Sociology data-Kernel. Unpublished https://doi.org/10.13140/RG.2.1.4971.4967
- David R, Feral J, Gachet S, Dias A, Blanpain C, Lecubin J, Diaconu C, Surace C, Gibert K (2015) SITIS 2015, 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). SITIS 2015, Bangkok, Thailand, nov. 2015. 2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS) https://doi.org/10.1109/sitis.2015.119
- Muñoz VM, Cohen-Nabeiro A, David R, Ivars Camáñez VJ, Nonell-Canals A, Senar MA, Couvet D, Feral J, Delavaud A, Tatoni T (2017) complexis conference proceding. complexis conference. Proceedings of the 2nd International Conference on Complexity, Future Information Systems and Risk https://doi.org/10.5220/0006379701440151